DOCUMENT RESUME

ED 257 075 CS 208 875

AUTHOR Kroll, Barbara
TITLE What Does Time Buy? Syntactic Accuracy and Discourse Fluency in ESL Composition.
PUB DATE [82]
NOTE 31p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Coherence; *Discourse Analysis; *English (Second Language); Higher Education; Second Language Instruction; *Second Language Learning; *Syntax; *Time Factors (Learning); *Writing (Composition); Writing Research
IDENTIFIERS *Freshman Composition

ABSTRACT

In comparing 50 English as a second language (ESL) compositions written in class to 50 ESL compositions written at home, accuracy on the syntactic level and fluency on the discourse or rhetorical level for the home and class conditions were examined. The subjects, twenty-five undergraduate foreign students enrolled in special sections of freshman composition for international students at the University of Southern California (USC), were randomly drawn in a stratified random sample to represent the five largest foreign groups at USC. Each contributed four essays to the data base. Two were written in class and two were written at home. Syntactic accuracy was measured using a ratio of words per error. The measure used to evaluate discourse fluency was a holistic score designed to measure adherence to organization and coherence only. While tests showed no statistical significance to the differences in class and home preformance, many of the subjects did show improved performance at home on an individual basis. The comparison between scores achieved for syntactic accuracy and discourse fluency shows no relationship between the two scores on individual compositions. The report concludes that time does not buy much for students in the improvement of either their syntax or their organization, and that the level of performance in these two areas is not interdependent. (DF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED257075

What Does Time Buy? Syntactic Accuracy and Discourse Fluency in ESL Composition

Dr. Barbara Kroll

Mailing Address:
933 S. Curson Avenue
Los Angeles, CA 90036
(213) 936-8288

Position:
Visiting Lecturer,
Dept. of Linguistics
University of California
Santa Barbara, CA 93106
(805) 961-3776

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Barbara Kroll
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS208875

Abstract

This study compares 50 ESL compositions written in class to 50 ESL compositions written at home. It further compares accuracy on the syntactic level and fluency on the discourse or rhetorical level for the class and home conditions.

To evaluate a data base of 100 essays, measures were developed to assess syntactic and discourse properties separately. Syntactic accuracy was measured using a ratio of words per error. The measure used to evaluate discourse fluency was a holistic score designed to measure adherence to organization and coherence only.

While T-tests showed that there was no statistical significance to the differences in class and home performance, many of the subjects did show improved performance at home on an individual basis. Furthermore, the comparison between the scores achieved for syntactic accuracy and discourse fluency shows no relationship between the two scores on individual compositions.

We can conclude that time does not buy very much for students in the improvement of either their syntax or their organization, and that level of performance in these two areas is not interdependent.

Writing is frequently the most difficult skill area for any language user, which is to say that writing is a challenging task in one's own language as well as in a second language. This difficulty does not exempt foreign students at American universities from having to write papers as part of their student careers regardless of their major subjects at school. However, it seems fair to say that this is a particularly difficult task for them. English as a second language (ESL) students must learn to create written products which illustrate mastery not only in all of the areas related to the rhetorical presentation of ideas, but also in the area of syntax, a Herculean task given the possibilities for error. For teachers to structure courses that will help foster progress in their ESL students, I believe that more research is needed on what the writing of non-native speakers actually looks like so that realistic goals can be established. We need to know what students do do in order to know what we can ask them to do.

One of the difficulties in establishing clear goals is the fact that the question of native-speaker proficiency is hardly a simple issue. There is just no written standard that can be said to represent the "ideal" written product in English. Therefore, we can't easily establish procedures for evaluating ESL writing in terms of adherence to some model of native-speaker writing. Furthermore, there is much documentation of the output of a type of non-successful native speaker often referred to as a "basic writer." In a now classic major study of a population of basic writers, Shaughnessy (1977) has indicated that mistakes learners make are often neither attempts to deliberately sabotage language in reckless disregard of its rules nor necessarily careless inattention to details. Rather, as her study shows, the mistakes

of learners are often the result of internally consistent and carefully worked out but misguided interpretations of language. Many errors on the part of ESL students stem from similar sources, and these errors often seriously interfere with the communicative efficacy of a piece of prose.

In this study, I would like to examine both the occurrence of error and to assess whether or not there is a connection between the level of syntactic accuracy and the overall organizational success of a student essay. Essentially what we are looking at is the question of whether it is possible to write a good essay in bad English or a bad essay in good English.

These issues will be addressed in terms of a research focus that has not received much attention previously, namely the issue of time. Does time buy a reduction of error and an improvement in manner of presentation? The major research focus of this study will thus be on how the factor of the time allowed for the preparation of an essay affects its success both on the syntactic and the discourse level.

Methodology

In this report, I will present a descriptive analysis of 100 essays written by 25 advanced ESL students. Much previous research on writing has analyzed compositions which were produced under strictly controlled conditions, with specified time limitations. In addition to analyzing compositions collected in the usual test/pressure situation, I will also be analyzing essays produced at a more leisurely pace, with more opportunity for the writer to think over the subject matter, check on grammatical principles, revise passages, and engage in whatever other behaviors can accompany the writing process over time that tend to disappear with a restricted time frame.

The compositions to be analyzed here are divided into two major groups: those produced in class and under pressure of time and those produced at home with 10-14 days preparation time. In the analysis of the essays, I will constantly be comparing the class essays to the home essays because we would like to know whether the removal of the time pressure has any effect on the actual products of the writing process. The description of the essays will first identify what it is that students produce and then what it is that time buys for them.

What I will show is that while the time allowed for the preparation of an essay often contributes to improvement for a writer both on the syntactic level and the rhetorical level, it does not necessarily create a sufficiently better essay to approach statistical significance. In my analysis, I will also discuss the relationship between grammatical accuracy and organizational success to show that the two occur independently of each other.

Hypotheses

Behind the close analysis of these 100 papers are the following three hypotheses:

(1) Writing produced in class under pressure of time will exhibit less control over syntax than writing produced outside of class (at home).

(2) Writing produced in class under pressure of time will exhibit a lower level of organizational skill than writing produced outside of class (at home).

(3) There is no necessary relationship between syntactic accuracy and discourse fluency.

The first hypothesis stems from Krashen's monitor model (Krashen, 1977;

1981), which predicts that given certain conditions, the learner can apply consciously learned grammar rules to alter and improve their written or spoken utterances. Briefly, monitor theory begins with the tenet that language is rule-governed behavior, but there are two ways in which this behavior becomes part of any one individual's performance. The first way is through acquisition, which Krashen (1977) defines as the internalization of linguistic abilities without conscious focusing on linguistic forms. The second path to attaining a linguistic skill is, in contrast to acquisition, a conscious process which Krashen (1977) calls learning. Krashen claims that conscious learning is available to the performer only in the role of a "monitor," which is the deliberate application of formal knowledge "to alter the output of the acquired system ... to improve accuracy" (Krashen, 1981, p.2). The monitor model may also account for the ability to edit a piece of writing, which could improve the overall presentation as well as the grammar. In this way, the monitor model relates to the second hypothesis as well.

An additional reason for comparing class to home compositions, besides an examination of whether the additional time allows the monitor to surface, is to respond to the inherent sense that writing in class is a very unnatural situation and perhaps cannot lead to work which is truly reflective of anyone's best capabilities. This philosophy is advanced in a report by Sanders and Littlefield (1975) who point out: "Unfortunately, the rigidly controlled essay test situation surely represents the ultimate in an artificial writing situation; as such, it is exactly the kind of situation shunned in many modern composition courses" (p.147). All in all, our intuitive sense is that writing produced with a 14-day deadline should be superior to writing produced with a 50-minute deadline.

The third hypothesis, relating to the relationship between the syntactic level and the discourse level, stems from the assumption that different skill areas contribute independently to the total "success" of a piece of writing. The multiplicity of skills involved further contribute to the overall difficulty of writing. Collins and Gentner (1980) make the following observation:

> Much of the difficulty of writing stems from the large number of constraints that must be satisfied at the same time. In expressing an idea the writer must consider at least four structural levels: overall text structure, paragraph structure, sentence structure (syntax), and word structure.... Clearly the attempt to coordinate all these requirements is a staggering job. (p.67)

If various aspects of writing are seen as levels which must be coordinated, it is reasonable to hypothesize that success may vary from level to level.

Subjects

The subjects in this study were all undergraduate foreign students enrolled in special sections of freshman composition for international students at the University of Southern California (USC), the highest level required composition course at the time. Twenty-five students were selected -- in groups of five -- to represent the five largest foreign language groups at USC: Arabic, Chinese, Japanese, Persian, and Spanish. The students were randomly drawn in a stratified random sample from five different sections of the course. Collectively, the subjects are felt to match the population of a typical class of 25 foreign students where students in any one class are heterogeneous both in their native language and in their language backgrounds.

Procedures

Each of the 25 subjects contributed four essays to the data base. Two of the essays were written in class in a test-like situation where the topics were not announced in advance. The subjects had to plan, write, and rewrite their essays within the space of one class lesson, typically 50 minutes. Two of the essays were written at home, and topics were distributed 10-14 days before the essay had to be handed in.

The findings to be reported here stem from two rather separate ways in which the data was coded. In the first procedure, to test Hypothesis 1, each composition was examined in detail and every single syntactic error on the sentence level was identified and labelled, except for spelling, which was overlooked (the same procedure followed by Neilson, 1979). The second procedure, to test Hypothesis 2, was to evaluate the essays holistically for adherence to principles of organization and coherence. In the holistic scoring, the essays were read and evaluated without regard to their syntactic accuracy. That is, the readers overlooked all syntactic errors in arriving at a holistic "discourse" score. To test Hypothesis 3, the holistic score can be compared to a score representing the syntactic accuracy of the same composition in order to determine whether or not there was any correlation between the level of success in each skill area. This last procedure provides answers to the question of whether it is possible to write a good essay in bad English or a bad essay in good English.

Syntactic Analysis

Because writing is not syntax alone and because the analysis of grammatical accuracy says little or nothing about the composition as a whole, I would

like to state two reasons why we must look at accuracy in writing. In the first place, grammatical accuracy is a major feature of standard written English. Secondly, one is drawn toward grammatical analysis because it is possible to codify, measure, and count various features within the grammar of the language, thus allowing the presentation of research findings in a standard format. The measurement of grammatical accuracy allows us to compare achievement from one essay to the next.

In closely examining each sentence in the corpus of essays, the criterion for deciding whether or not an error had been committed and, if so, what type of error, was to determine what "syntactic reconstruction" could most easily and economically render the sentence into acceptable English given the context. Following this procedure, a total of 33 different categories of error were identified for coding. In all, a total of 2,307 errors were identified and labelled in the corpus of 100 compositions totalling 1,599 sentences (28,444 words).

Distribution of Errors

The 33 categories of error can be grouped into four major classes, viz. sentence structure errors, verb errors, reference errors, and word-level errors; and two unique classes, viz. article errors and punctuation errors. The distribution of errors into these six categories is shown in Table 1, which tabulates the percentage of each category that was found in the 50 class compositions and the 50 home compositions. The percentages in each column total 100% (± rounding error), as it is not possible for there to be a reduction of percentage in one column or the other. In each case, 100% of the errors are being divided into the corresponding categories.

Insert Table 1 about here

In some sense, the first class of error, sentence structure, leads to more serious violations of English, for in general, errors found in this class sometimes would require a total reworking of the sentence before meaning could be accurately derived. In contrast, errors of word-level choice, which occurred more frequently, rarely interfered with communication, and the intended meaning could almost always be readily grasped despite the deviation from standard written English.

As the writers in this study were all advanced learners of English, it should not seem surprising that a much smaller percentage of the error corpus fell into the categories grouped under sentence structure as compared with those categories grouped under word-level choices. When errors did occur, they tended not to be in the class of the more serious violation.

Another way to assess the difference between the occurrence of sentence structure and word-level errors is to consider the rank order distribution of the 33 categories. Such a calculation indicates that the average of the ranks for all 14 sentence structure errors is 20.8, while the average of ranks for the 7 word-level errors is 11.1. This reinforces the finding that on the performance level, subjects fared worse with choices on the word level than with choices related to structures beyond the word. In correlation tests of the distribution of errors in class and at home, the Spearman rho for all 33 errors is .904, with a significance level of $p < .05$. As the very high correlation shows, there is a very similar distribution of the error categories in

class and home compositions. We can only point to tendencies in differences between performance on class and home compositions; we cannot establish that the students' error pattern was significantly different. (For a full discussion of the categories of error and their rank order, see Kroll, 1982.)

Accuracy Ratio

Apart from tabulating the total number of errors and categorizing their distribution, we need to measure the occurrence of error within the framework of the total composition. That is, we need to measure the relationship between error (what is wrong) and accuracy (what is right). The importance of this step can be seen by taking a hypothetical example. Suppose Student A made 10 errors of any nature in a composition of 10 pages totalling 2,500 words while Student B made 10 errors of a similar nature in a composition totalling 150 words. We could not call their performances similar even though the quantity of errors is similar. The counting of syntactic errors only becomes meaningful within a consideration of the range of opportunity for error.

In the present study, the sentence was taken as the basic unit under focus in checking the syntax. Each word within the sentence was seen as a possible "opportunity" for error as well as the sentence itself (the latter giving rise to such categories as run-on and fragment). Using the total number of words in a composition and tabulating the number of errors is one of the standard measures used in forming the basis for a kind of accuracy ratio (see, for example, Brière, 1966; Garnes, 1978).

In Table 2, a general or overall accuracy ratio is shown for each language group. These figures are derived from a two-step procedure. First, the total number of words in each composition is divided by the number of errors

in each composition. Then the resulting individual accuracy scores for the subjects in each language group are averaged for class and home compositions.

Insert Table 2 about here

The numbers in Table 2 can be read, for example, as follows: there are an average of 19.4 words between each error in class compositions written by Arabic speakers. The higher the number, the fewer errors are found proportionately, while conversely, the lower the number, the greater the proportion of errors. The highest number here, 22.7, shows the Arabic speakers had one error every 22.7 words in the home corpus in contrast to the lowest number, 7.8, representing one error every 7.8 words for the Japanese class corpus. Since these numbers are ratios, they can be compared, and we can say, for example, the Japanese writing in class had more than twice as many errors as the Arabs writing at home. Taken as a whole for both class and home, the group averages show the Arabic speakers as having the most accurate prose and the Japanese the most flawed.

Also according to this table, we see that all of the five language groups show a proportionately better performance (at least minimally) in home compositions over class compositions, averaging out to means of 14.8 in class vs. 18.0 at home. What accounts for this is the relation between the increase in the number of words written at home and the corresponding increase or decrease in the number of errors; all groups produced more words at home, but the proportion of errors went both up and down.

The Spanish group performed only marginally better at home than in class.

In their case, the number of words increased by 8.5% for home compositions, but the number of errors increased by 21.5%. In actual fact, the lack of marked improvement in the Spanish group's home performance mostly derives from problems with only two of the 33 error categories: word choice and word form, as summarized in the chart below:

Error Category	Class Error Count		Home Error Count	
Word choice	7.9%	N=24	13.4%	N=49
Word form	7.3%	N=22	10.4%	N=38
Total:	15.2%	N=46	23.8%	N=87

Although the Spanish-speaking subjects produced by far the most lengthy corpus of all the groups, their high incidence of error, particularly in the two categories cited above, gives them the second lowest accuracy ratios after the Japanese subjects, who collectively produced the shortest corpus. This latter fact underscores, perhaps, the poor performance of the Japanese whom we might say had a reduced opportunity for error given their shorter papers but who managed to make proportionately more errors than any other group.

Discussion of Syntactic Findings

Of the 33 categories, the one showing the greatest reduction in error rate moving from class to home was verb tense, while the one showing the greatest increase in error rate was article usage. Verb tense accounted for 7.1% of the class errors, which was the fourth ranking error. For home compositions, the percent of error was reduced to 4.5%, which ranked seventh of all errors made at home. Articles accounted for 10.8% of class errors, the third ranking error, but rose to 14.0% of the errors at home, becoming the number one ranking error.

We might surmise that is was possible for subjects to improve their per-

formance at the level of verb tense selection when they had sufficient time to allow them to proofread their papers. This would help them access their monitors, which could easily include a knowledge of the rules for verb tense. However, inasmuch as the rules for articles are complex and difficult, and not the kind of rules that ESL students can generally articulate, it is more likely that their performance would be haphazard, and extra time would not help them. In actual fact, all five language groups showed worse performance with articles at home, though this was especially true for the Persians, and only marginally true for the Arabic and Spanish speakers.

According to the accuracy ratios examined on a subject by subject basis (see Kroll, 1982), 64% of the subjects did better at home while 36% performed with greater syntactic accuracy in class. An additional tabulation of the average number of different categories of error shows that 72% of the subjects had a narrower range at home, while 24% had a narrower variety of error types when writing in class. The percentage of subjects who simultaneously showed a better accuracy ratio at home together with a narrower range of error categories was 60%, while 24% show both a better accuracy ratio and a narrower range of error in class. All in all, we would have to say that subjects did write better when they had more time to prepare their essays although their improvements were not statistically significant.

Holistic Evaluation

The second measure which was used to code the data was to assign a holistic score to represent each essay's adherence to principles of organization and coherence, or what might be termed the pure discourse features of an essay. The key to holistic evaluation is to establish some sort of written

guidelines or rubric which sets out the principles with which a reader is to judge a set of essays. These procedures of holistic evaluation are reviewed by Myers (1980) and Najimy (1981), while Cooper (1977) gives an overview of the varieties of holistic evaluation.

The key to the rubric which was developed was that readers had to overlook and ignore all errors not related to the features directly under examination, and to focus solely on the "larger" issues of discourse. This procedure necessitated reading through the errors of syntax and attending only to the level of organization and coherence. In other words, the scores that were assigned to each individual essay rated them from 1 to 6 on a package of isolated discourse properties and not on the basis of the essay as a whole. The essays were being scored as if they had no grammatical errors when in fact they averaged one error every 15 to 18 words. (Focusing beyond the level of syntax to evaluate essays was similar in intent to work done by Freedman, 1977, though quite different in procedure.)

For the rubric used in this study, those features which most contributed to a high holistic score included (but were not limited to):

1. focused limitation of the topic
2. remaining on the focused topic throughout the essay
3. effective use of paragraphing
4. consistency in point of view
5. logical sequencing of ideas
6. artful use of transitions

Identifiable features which lowered the score of a paper included (but were not limited to):

1. noticeable introduction of irrelevancies

2. failure to provide a clear sense of purpose
3. shifting point of view
4. infelicitous, inappropriate or non-existent transitions
5. assumption of an argument's validity with no development of the argument

Two ESL teachers experienced in the reading of compositions and with previous practice in holistic grading were trained as graders for the essays. After the training session, readers went through the essays which had no markings to indicate whether they had been written in class or at home, and also had no markings as to the language origin of the subject. The readers were able to achieve an inter-rater reliability coefficient of .85, showing a high degree of agreement on the scores.

Holistic Performance by Language Groups

In order for a paper to merit an upper half grade (4, 5, or 6), it had to demonstrate control over both the structure of the essay and the structure of individual paragraphs. With this in mind, even a score of 4 demonstrates a fair degree of discourse fluency. Table 3 summarizes the percentage of upper half and lower half scores for each language group in class and at home. (Since the number of compositions in each cell is 10, the percentage corresponds to the number of compositions, i.e., 50% equals five compositions.)

Insert Table 3 about here

As seen in Table 3, for class compositions, the Chinese, Persian and Spanish subjects had the same distribution of scores (40% upper vs. 60% lower) with the Arab breakdown for the scores fairly similar at 50% and 50%.

Only the Japanese subjects had a noticeable difference in upper half and lower half scores. The preponderance of lower half scores for the Japanese both at home and in class ranks the Japanese essays as the worst in terms of discourse fluency. At the opposite extreme, the Arabs show the largest percentage of scores in the upper half and emerge as the group with the best discourse fluency.

The actual mean scores of all language groups is shown on Table 4. All five language groups averaged slightly higher scores for the home compositions (including the Japanese), but T-tests for significance indicate that none of the differences are statistically significant. However, we should not discount the evidence of improvement in the home condition.

Insert Table 4 about here

Distribution of Holistic Scores

We can also consider the distribution of the scores in the corpus irrespective of the subjects' language background. This is done by tabulating the total number of essays which received each of the possible scores. Such a breakdown is shown in Table 5 (as computed in percentages). Not surprisingly, the majority of the scores cluster around the mid-range of possible scores. Here we see that of class compositions, 56% received scores ranging from 3 - 4.5, while 52% of the home compositions did, indicating that just over half of the papers were neither very poorly organized nor very well organized.

Insert Table 5 about here

The percentage of scores at the very bottom of the scale decreased by half from class to home (from 16% to 8%), while the percentage of highest scores tripled (from 2% to 6%). Even though the actual number of "6" scores is quite small, it is important to note that more such scores appeared in home compositions. Remember that the readers assigned scores to papers without knowing whether they had been written in class or at home.

Totalling the percentage for lower half scores vs. upper half scores shows that while lower half scores clearly outnumbered upper half scores in class compositions (62% vs. 38%), the gap was considerably narrowed in home compositions (52% vs. 48%). In other words, there was a shift of 10 percentage points in the movement from class to home in the direction of more upper half scores. Another way to point to the difference is by noting that the ratio of bottom to top scores in class was slightly more than 2:1, while the ratio of bottom to top scores at home was approximately 6:5. All of this shows that in group terms, the subjects wrote better organized compositions at home. With increased time, they were able to reduce somewhat the incidence of the poorest level of performance and noticeably increase the incidence of above par performance.

Discussion of the Holistic Results

In terms of home vs. class essays, it appears that there is a tendency for subjects to write higher rated essays at home, based on the group mean scores of 3.2 for class vs. 3.6 for home. In fact, at least minimally, all five groups scored higher means for home essays than for class essays, though none of these differences are statistically significant. On an individual basis, 52% of the subjects (N=13) averaged higher scores for their two home essays over their two class essays, while only 32% of the subjects (N=8)

averaged lower scores. In his dissertation, Hartvigsen (1981) found a similar distribution of holistic scores for in-class vs. out-of-class essays written by native speakers. In his sample, approximately 50% of the subjects had higher mean scores for their out-of-class essays, while only 14% had higher means for in-class essays (with 36% of his subjects having the same means). Despite the lack of statistical significance of the findings in the present study, we would have to say that there is some support for Hypothesis 2, namely, students can produce better organized essays given more time.

The Interface of Syntactic and Rhetorical Assessments

Once we have been able to establish scores which assess the syntactic accuracy of each composition and assign it an evaluation of discourse/rhetorical effectiveness, it is now possible to address the third hypothesis of the study which predicts that there is no necessary relationship between these two areas.

If we consider the holistic score of the individual compositions together with their accuracy ratios, we can see that there is no real pattern to the way the two scores occur as pairs in the 100 compositions (see Kroll, 1982 for the actual figures).

In fact, the results of the Spearman correlation test for these scores show that the two scores for the compositions are not statistically correlated. The value of rho (the correlation coefficient) in each case is exceedingly low: rho = .083 for class essays and rho = .043 for home essays. Neither of these values is significant at the $p < .05$ level. If the scores were correlated, that would show that one score is connected to the other and potentially predictive of the other. In other words, such a correlation would

show that there is a relationship between discourse fluency and syntactic accuracy. However, as that is not the case, we can claim rather that the data support Hypothesis 3, viz., there is no necessary relationship between syntactic accuracy and discourse fluency.

The lack of relationship between the words/error ratios and the holistic scores is also revealed in tabulating the ranges of accuracy ratios that co-occur with each possible holistic score. This is shown in Table 6.

Insert Table 6 about here

Both the lowest holistic scores of 1 - 2.5 and the highest holistic scores of 5 - 6 co-occur with relatively high and relatively low words/error ratios. If there were a correlation to the scores, we would expect to find low scores on the holistic scale corresponding to low scores in the range of accuracy ratios, and so on.

These results can provide answers to the questions posed earlier in the study of whether it is possible to write good essays in bad English and bad essays in good English.

In fact, the answers to both these questions is the same. Subjects can show control over the level of either syntax or rhetoric while simultaneously showing poor control at the other level. So, we cannot predict their ability to perform in one area on the basis of their performance in the other one. We do not teach one skill by teaching the other.

Conclusions and Implications

In this study, we have seen a definite but statistically insigniciant tendency for students to write better at home than in class. Based on the measures used in this study to assess syntactic accuracy and discourse fluency, it appears that students improved at both levels at home. In fact, 64% of the subjects averaged higher syntactic accuracy ratios when writing at home while 52% of the subjects averaged higher holistic scores for their home compositions. Although more students improved at the syntactic level, the gains made at the rhetorical level were larger, that is, the increase in the percentage of occurrence for upper half holistic scores in the home compositions appears to be more noticeable than the increase made in the accuracy ratios reflecting an increased number of words between errors.

Intuitively, we would expect that having additional time to prepare a writing assignment would allow a student additional time to think through an interpretation of an assignment before actually beginning to write. This gestation period may contribute more to allowing an organizational strategy to surface than it can contribute to the emergence of correct grammatical forms. In effect, the pre-writing or pre-thinking of an essay is almost invariably on the organizational level. Writers attend more to content than to form before sitting down to write. Attention to the grammatical level is almost always part of the revision process, where re-reading allows one to check rules and attend to form, i.e., monitor the output. The fact that the syntactic level improved only slightly in moving from class to home (although it did improve for all groups) may attest to the fact that subjects did not spend much, if any, time in revision at all. Without a specific injunction to

focus on form, perhaps few students do. Perhaps so few rules can be learned and made a part of conscious knowledge that it is really only possible to monitor a very small percentage of the language output. In any case, surely many teachers have had the suspicion that some essays handed in after 10-14 days appear to have been written in the same 50 minutes allowed in one class period, if not written in even less time than that.

Many students appear to be convinced that they cannot write under pressure of time and that if they were allowed to prepare essays outside of class, their performance (and presumably their grades) would noticeably improve. Based on the results of this study, this is only marginally true. Furthermore, the distribution of the specific language errors is remarkably similar in essays written in class and at home. Therefore, it is fair to say that in-class writing samples are as representative of a student's grammatical abilities as out-of-class samples.

Perhaps the students who wrote better in class or those who made only slight gains did not know enough about what constitutes good writing in the first place, so that their performance was accidental rather than calculated. They may have had no way to access information about how to write better because they did not have any mental formulation of what constitutes good writing. In such cases, time could not buy students anything because they would not know how to proceed in the task of writing. Such students may attack every task with the same lack of skill regardless of the conditions they are writing under.

References

Brière, Eugene J. (1966). Quantity before quality in second language composition. Language Learning, 16, 141-151.

Collins, Allan and Gentner, Dedre. (1980). A framework for a cognitive theory of writing. In L.W. Gregg & E.R. Steinberg (Eds.), Cognitive processes in writing (pp.51-72). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, Charles R. (1977). Holistic evaluation of writing. In C.R. Cooper & L. Odell (Eds.), Evaluating writing: Describing, measuring, judging (pp. 3-32). Urbana, IL: National Council of Teachers of English.

Freedman, Sarah W. (1977). Influences on the evaluators of student writing. Unpublished doctoral dissertation, Stanford University.

Garnes, Sara. (1978, August). Error reduction in remedial composition at the college level. Paper presented at the Fifth International Congress of Applied Linguistics, Montreal, Canada.

Hartvigsen, M. Kip. (1981). A comparative study of quality and syntactic maturity between in-class and out-of-class writing samples of freshmen at Washington State University. Unpublished doctoral dissertation, Washington State University.

Krashen, Stephen D. (1977). The monitor model for adult second language performance. In M.Burt, H. Dulay & M. Finnocchiaro (Eds.), Viewpoints on English as a second language (pp.152-161). New York: Regents.

Krashen, Stephen D. (1981). Second language acquisition and second language learning. Oxford: Pergamon Press.

Kroll, Barbara. (1982). Levels of error in ESL composition. Unpublished doctoral dissertation, University of Southern California.

Myers, Miles. (1980). A procedure for writing assessment and holistic scoring. Urbana, IL: National Council of Teachers of English.

Najimy, Norman C. (Ed.) (1981). Measure for measure: A guidebook for evaluating students' expository writing. Urbana, IL: National Council of Teachers of English.

Neilson, Brooke. (1979). Writing as a second language: Psycholinguistic processes in composition. Unpublished doctoral dissertation, University of California at San Diego.

Sanders, Sara E. and Littlefield, John H. (1975). Perhaps test essays can reflect significant improvement in freshman composition. Research in the Teaching of English, 9, 145-153.

Shaughnessy, Mina. (1977). Errors and expectations. New York: Oxford University Press.

Table 1

Distribution of Major Classes of Syntactic Error

Error Category	Class Percentages	Home Percentages
Sentence Structure Errors	21.3	18.0
Verb-Centered Errors	16.8	12.3
Reference Errors	6.9	5.9
Word-Level Choices	32.6	36.6
Articles	10.8	14.0
Punctuation	11.8	13.5
	N= 1142	1165

Table 2

Mean Individual Accuracy Ratios by Language Group

	Class	Home
Arabic	19.4	22.7
Chinese	15.9	20.7
Japanese	7.8	12.0
Persian	16.8	19.3
Spanish	14.0	15.5
Group Mean:	14.8	18.0

Table 3

Upper Half vs. Lower Half Scores by Language Group

	Class		Home	
	Upper Half	Lower Half	Upper Half	Lower Half
Arabic	50%	50%	80%	20%
Chinese	40%	60%	60%	40%
Japanese	20%	80%	10%	90%
Persian	40%	60%	30%	70%
Spanish	40%	60%	60%	40%
Group Means	38%	62%	48%	52%

Table 4

Holistic Scores by Language Group

	Class Mean Score	Home Mean Score
Arabic	3.8	4.2
Chinese	3.4	3.7
Japanese	2.4	2.8
Persian	3.2	3.3
Spanish	3.3	4.1
Group Mean	3.2	3.6

Table 5

Distribution of Holistic Scores

Score	Class	Home
1, 1.5	16%	8%
2, 2.5	14%	18%
3	32%	26%
4, 4.5	24%	26%
5, 5.5	12%	16%
6	2%	6%
Total lower half =	62%	52%
Total upper half =	38%	48%

Table 6

Range of Words/Error Ratios Found for Each Holistic Score

	Class	Home	Class & Home
Holistic Score	Range Words/Error	Range Words/Error	Range Words/Error
1, 1.5	5.3 - 35.0	7.0 - 21.4	5.3 - 35.0
2, 2.5	9.3 - 32.2	7.4 - 54.0	7.4 - 54.0
3	5.0 - 17.6	6.0 - 18.5	5.0 - 18.5
4, 4.5	5.4 - 33.4	6.7 - 86.3	5.4 - 86.3
5, 5.5	8.8 - 66.4	10.0 - 37.8	8.8 - 66.4
6	- 22.3	9.1 - 28.4	9.1 - 28.4